DOCUMENT RESUME

ED 316 553 TM 014 497

AUTHOR Robertson, Gary J.
TITLE Current Status of the Test User Qualification Process.
PUB DATE Aug 89
NOTE 14p.; Paper presented at the Annual Meeting of the American Psychological Association (97th, New Orleans, LA, August 11-15, 1989).
PUB TYPE Reports - Research/Technical (143) -- Speeches/Conference Papers (150) -- Tests/Evaluation Instruments (160)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Examiners; *Experimenter Characteristics; Measurement Techniques; *Qualifications; *Researchers; Research Methodology; Testing Problems; *Test Use
IDENTIFIERS *Screening Procedures; Test Publishers; Test User Qualifications Working Group; *User Characteristics

ABSTRACT

The development of procedures by which test publishers can screen individual purchasers to minimize the possibility of test misuse is described. The Test User Qualifications Working Group (TUQWoG), an outgrowth of the Joint Committee on Testing Practices, was formed to address the prevention of test misuse through improved methods of screening test purchasers. In defining the scope of its work, the group decided to study commercially published tests, excluding licensure and certification examinations, college entrance tests, and entrance tests for professional schools. The studies conducted by TUQWoG focused on the individual test purchaser, not the institutional purchaser. Study 1 identified the content domain of test user competency through a critical incidents study with 62 experts. Study 2 examined the key test user behaviors for tests in general by having 19 experts rate the 86 key test user behaviors obtained in Study 1. Study 3 described and evaluated common factors of test misuse through the input of 200 experts who rated 76 tests. Study 4 encompassed the development of actual test purchaser forms using data from the first three studies. Two test publishers developed these forms, and a sample form is provided. A technical report issued by TUQWoG contains guidelines for developing a purchaser qualification form. Two tables and one figure illustrate the process. (SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED316553

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

GARY J. ROBERTSON

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

CURRENT STATUS OF THE TEST USER QUALIFICATION PROCESS

Gary J. Robertson, Ph.D.
American Guidance Service
Circle Pines, MN 55014

A paper presented in August, 1989, as part of a symposium at the
American Psychological Association Convention
New Orleans, Louisiana

TM014497

Current Status of the Test User Qualification Process

The focus of this discussion will be the test purchaser: the individual who places an order for standardized testing materials with a publisher. The institutional purchaser of tests, while unarguably important, is not a focal point of this paper because it was necessary to place limitations on the scope of the work described here.

The "test user qualification process" used in the title of this paper refers to the procedures by which test publishers screen individual test purchasers to minimize the possibility of test misuse. There have been concerns for many years about tests reaching only those individuals with appropriate professional training. One of the earliest evidences of this concern was the publication by APA of the first set of ethical standards for the sale and distribution of psychological tests (APA Committee for Ethical Standards for Psychology, 1950). In 1950, the Committee released its Ethical Standards for the Distribution of Psychological Tests which consisted of 18 principles governing the sale and distribution of psychological tests. This code of ethics was developed over a four year period and was based on critical incidents dealing with ethical behavior. The code was formally adopted by APA in 1953 (Golann, 1970).

Other evidences of concern by professional societies about the proper distribution and use of tests are shown in the various editions of the technical standards for the development of tests and the codes of ethics developed by such organizations as APA, the American Association of Counseling and Development (AACD), the American Speech-Language-Hearing Association (ASHA), and the National Association of School Psychologists (NASP), to mention but a few.

Concern about the proper use of tests was a major topic at a 1984 conference of test publishers held at the APA headquarters in Washington, DC. An outgrowth of this meeting was the establishment in 1985 of the Joint Committee on Testing Practices (JCTP), originally sponsored the American Educational Research Association and the National Council on Measurement in Education, in addition to APA. The stated mission of JCTP is "To promote quality assurance in testing by working with test publishers and with major test user groups," (Eyde, et al., 1988, p. 17). The Test User Qualifications Working Group (TUQWoG) was formed to address the problem of the prevention of test misuse through

improved methods of screening test purchasers (Eyde, et al, 1988, p.11). The research base established by TUQWoG and the various work products based on those data reflect the current status of efforts to improve test purchaser screening procedures. A description of the work of TUQWoG constitutes the remainder of this paper.

Purpose of TUQWoG

TUQWoG, which was founded in February 1985, adopted the following mission statement:

> Consistent with professional standards and ethical principles, to establish a procedure for classifying tests that will lead to a uniform system for assessing test user qualifications (Eyde, et al., 1988, p. 23).

Furthermore, it was implicit in the work of the group that any undertaking or work product must, of necessity, be consistent with extant ethical standards and principles, including the 1985 *Standards for Educational and Psychological Testing*. An added precept was that test user qualifications should identify the specific training, certification, or experience needed in order to use tests properly.

The specific objectives underlying the work of TUQWoG were as follows:

1) To facilitate access by qualified users to testing materials.

2) To facilitate and improve test publishers screening procedures.

3) To protect test takers from test users who administer and interpret tests they are not qualified to use.

4) To improve the education of those who use tests.

Characteristics Desired in TUQWoG Work Products

The work of TUQWoG was based on several carefully formulated principles that guided the group's work and shaped the final work products. The TUQWoG members decided at the beginning of their work that a scientifically based model test user qualification

system should be developed; moreover, the system should emphasize test user behaviors, not job merely titles. Specifically, test user competencies were seen as the focal point of any test user qualification system. The practical constraints of the business setting in which test user qualification forms are used dictated that any forms developed should be as streamlined and objective as possible in order to increase the likelihood of their adoption and use by test publishers. The group also felt that its mission was broader than just forms development and saw itself in the role of educating test users and potential users about proper test use. This broader orientation became apparent after the group had collected a vast array of data on test user behavior, as outlined in later sections of this paper. From the outset, TUQWoG adopted an interdisciplinary focus for its activities and established liaison relationships with some 20 different professional organizations that contributed significantly toward achieving the TUQWoG objectives (Eyde, et al., 1988).

General Characteristics of TUQWoG Research

In delimiting the scope of its work, TUQWoG decided to focus on commercially published tests in its studies of test user behavior. Specifically excluded were certification and licensure examinations, college entrance tests, and professional schools admissions tests. Of central concern were the so-called "off-the-shelf" tests sold by commercial test publishers who fill orders by mail or telephone. Furthermore, the focus from the standpoint of end use of the tests was the individual test purchaser, not the institutional purchaser who typically orders tests on institution letterhead. Practical constraints precluded studying institutional purchasers.

The major source document for establishing the tests to be used in TUQWoG research was Anastasi's listing in her 1982 text (pp. 669-682). This listing is comprehensive and includes the most widely used tests of various types. Test use was felt to be a job function in that it requires knowledges, skills, behaviors, and credentials common to the use of a variety of tests in myriad settings. For this reason, two systematic job analysis methods were used to identify test user competencies: Flanagan's (1954, pp. 327-358) critical incident method and Primoff's (1975) Job Element Method. These methodologies were used as the basis for the first two of the four studies described below.

Specific TUQWoG Research Studies

The different aspects of TUQWoG's work can best be described in terms of four distinct phases. These are considered research studies and are summarized briefly below. The results of these studies formed the basis for TUQWoG's work products.

Study 1. The purpose of this study was to identify the content domain of test user competency. Critical incidents of actual test misuse were gathered from experts such as test reviewers, test and textbook authors, and instructors of various types of assessment courses. A total of 62 experts participated in the critical incident study; 47 different tests were mentioned in the critical incidents. Primoff used his Job Element Method to abstract 86 key test user behaviors which formed the basis for further research(Eyde, et al., 1988, p. 25).

The following example shows a critical incident submitted by one of the experts contacted:

Errors in scoring subtests (careless as well as systematic bias such as "halo"; clerical errors in computing subtest scores and global IQs.)

From this critical incident Primoff abstracted the following key tests user behavior (one of 86 subelements of proper test use):

---Avoiding errors in scoring and recording.

Other examples of the 86 subelements of proper test use are:

---Giving standard directions as prescribed.

---Considering errors of measurement in the interpretation of tests.

---Knowing that multiple sources of information that agree are needed before proper inferences can be drawn from test results.

The entire list of 86 subelements of proper test use may be reviewed by consulting Appendix I of the TUQWoG Technical Report (1988, pp. 102-105).

Study 2. The purpose of Study 2 was to determine the key test user behaviors for tests in general. This determination was made

by having 19 experts rate the 86 key test user behaviors obtained in Study 1 on the following four dimensions:

a) **Trouble Likely**--how much trouble would result from NOT considering this subelement in qualifying test users?

b) **Barely Acceptable**--is this subelement a characteristic of test users who are marginally (barely) acceptable?

c) **Superior**--is this subelement a trait of a SUPERIOR test user?

d) **Practical**--is this subelement a practical, feasible requirement of test users?

The results of Study 2 were important in establishing the breadth or comprehensiveness of the 86 key subelements of proper test use. This degree of importance was determined by computing the Total Value, a Job Element statistic, for each of the 86 test user behaviors (Primoff, 1975). Two basic elements emerged: 1) Knowledge of the test and its limitations and 2) Acceptance of responsibility for proper test use. These were two important descriptors useful in classifying the subelements.

In addition to determining the degree of comprehensiveness of the 86 subelements, Study 2 also permitted the identification of the most basic or minimum requirements for competent test use. These 12 "screenouts" are so essential for proper test use that their absence on the part of a test user would seriously question that user's fitness for properly using any test. Examples of these "screenouts" are as follows:

a) Not making photocopies of copyrighted test materials

b) Keeping scoring keys and test materials secure

c) Not assuming that norms for one group automatically apply to other groups

The entire list of "screenouts" appears in the TUQWoG Technical Report (Eyde, et al., 1988, p. 84) and was of considerable importance in the development of the test purchaser qualification forms described in Study 4.

Study 3. The purpose of Study 3 was to describe and evaluate the common factors of test misuse for empirically-based clusters of

tests. In order to achieve this goal, it was necessary to have individual tests rated on the list of 86 key subelements of test misuse used by the raters in Study 2 to rate tests in general. In Study 3, however, specific tests were rated only on the **Trouble Likely** dimension in contrast to Study 2 where tests in general were rated on four dimensions.

More than 200 experts rated 76 tests that represented the various types of tests included in the listing in Anastasi's text. In all, 487 ratings were obtained. Six or more ratings were obtained for 50 of the tests rated and these constituted the data-base used in the analyses described below.

First, a principal factor analysis with varimax orthogonal rotation was used on the 487 **Trouble Likely** ratings for 50 tests. The result was the identification of the seven factors of proper test use shown in Table 1. Also shown in Table 1 is the subelement having the highest loading on that factor.

(Insert Table 1 here)

In order to obtain a practically useful system that would stand some hope of acceptance by test publishers, it was necessary to group tests into categories. The basis for accomplishing this was Ward's minimum variance cluster analysis (Romesburg, 1984), which was applied to the 487 ratings for 50 tests. The resulting taxonomy for the 50 tests with 6 or more raters is shown in Table 2.

(Insert Table 2 here)

Use of the information presented in Tables 1 and 2 greatly facilitates the identification of the importance of test misuse factors for each test cluster. It is also possible to identify the specific subelements that are of greatest importance for each test use factor/test cluster combination. This is the information of greatest interest to test publishers for test purchaser forms development.

Study 4. The purpose of Study 4 was the development of actual test purchaser forms using the data from the three studies undertaken by TUQWoG. Two test publishers have, at this time, developed such forms.*

* American Guidance Service and Consulting Psychologists Press.

Data from either Study 2 or Study 3 can be used, as required, to develop test purchaser forms. The classification structure for the 86 subelements that emerged from Study 2 was used by the writer to develop a new test purchaser form for use by American Guidance Service. This form is shown in Figure 3. Study 2 is useful for publishers that publish more than a single type of test. Study 3 will be more useful to publishers of a single type of test; i.e., objective personality inventories. Complete directions for developing test purchaser forms are provided in the TUQWoG Technical Report (Eyde, et al., 1988, pp. 43-50).

Discussion and Summary

The research process described in this paper is important because for the first time it is now possible to evaluate test purchaser qualifications by using data available from a large study of the factors of test misuse and their relationship to different types of tests. Such information has not been available previously. A study undertaken by the writer (Robertson, 1986) showed evidence of considerable variability among 13 major U.S. test publishers in information asked of persons applying to purchase tests. The TUQWoG database provides a way to achieve greater uniformity among publishers in the way in which potential test purchasers are screened for test purchase approval. In short, it provides a better match between the requirements for good test use and specific test purchaser characteristics than do existing procedures. The TUQWoG results are only now ready for wide dissemination to test publishers and other interested professionals. It is hoped that publishers will revise their current test purchaser screening procedures to take into account the information now being made available by TUQWoG. Such action should eventually translate into the improved use of tests which, in turn, will bring a number of needed benefits, including an improvement in experiences and services offered to examinees.

References

Anastasi, A. (1972). Psychological Testing (5th ed.). New York: Macmillan

Eyde, L. D., Moreland, K. L., Robertson, G. J., Primoff, E. S., & Most, R. B. (1988). Test user qualifications: A data-based approach to promoting good test use. Issues in Scientific Psychology. Washington, DC: American Psychological Association, Science Directorate.

Flanagan, J. C., (1954). The critical incident technique. Psychological Bulletin, 51, (4), 327-358.

Golann, S. E., (1970). Ethical standards for psychology: Development and revision, 1938-1968. Annals of the New York Academy of Sciences, 169 (2), 398-405.

Primoff, E. S. (1975). How to prepare and conduct job element examinations. U.S. Civil Service Commission Personnel Research and Development Center, Technical Study 75-1. Washington, DC: U.S. Government Printing Office (NTIS No. PB84-166- 743).

Robertson, G. J. (1986). Establishing test purchaser qualifications. In R. B. Most (Ed.), Test purchaser qualification: Present practice, professional needs, and a proposed system. Issues in Scientific Psychology. Washington, DC: American Psychological Association, Science Directorate.

Table 1
Factor Loadings for Behavior Illustrating Each Factor

Factor	Behaviors (Subelements)	Loading
1. Comprehensive Assessment: making appropriate interpretations	Follow-up to get facts from psychosocial history to integrate with test scores, as part of interpretation	.68
2. Proper Test Use	Acceptance of responsibility for competent use of the test	.62
3. Psychometric Knowledge	Consideration of standard error of measurement	.72
4. Maintaining Integrity of Test Results	Making clear that cut-off scores imposed for placement in special programs for the gifted are questionable because they disregard measurement error	.69
5. Accuracy of Scoring	Use checks on scoring accuracy	.76
6. Appropriate Use of Norms	Not assuming that a norm for a given job (or reference group) applies to a different job (or reference group)	.65
7. Willingness and Ability to Provide Interpretive Feedback to Clients	Willing and able to give interpretation and guidance to test takers, in counseling	.74

Table 2

Eight Clusters of 50 Commercial Tests Using Ward's Minimum Variance Clustering Method Based on Ratings of Test Misuse

Cluster A: Group educational tests	Cluster B: Ability & preference tests	Cluster C: Learning disability/ neuropsychological tests	Cluster D: Individual intelligence tests	Cluster E: Readiness tests	Cluster F: Objective personality tests	Cluster G:	Cluster H: Projective personality tests
California Achievement Test	Activity Vector Analysis	Bayley Scales of Infant Development	Full-Range Picture-Vocabulary Test	Boehm Test of Basic Concepts	Bem Sex Role Inventory	Self Directed Search	Rorschach
Cognitive Abilities Test	Armed Services Vocational Aptitude Battery	Bender Visual-Motor Gestalt Test	Goldman-Fristoe-Woodcock Auditory Skills	Metropolitan Readiness Tests	California Psychological Inventory		Rotter Incomplete Sentences Blank
Iowa Tests of Basic Skills	Bennett Mechanical Comprehension Test	Denver Developmental Screening Test	Illinois Test of Psycho-linguistic Abilities	Peabody Individual Achievement Test	MMPI		Thematic Apperception Test
Iowa Test of Educational Development	Differential Aptitude Tests	Goldman-Fristoe Test of Articulation	Kaufman Assessment Battery for Children	Slosson Oral Reading Test	Rotter Internal-External Locus of Control Scale		
Metropolitan Achievement Tests	Kuder Occupational Interest Survey	Halstead-Reitan Neuropsychological Battery	Peabody Picture Vocabulary Test	Stanford Diagnostic Reading Test	State-Trait Anxiety Inventory		
Otis-Lennon School Ability Test	Myers-Briggs Type Indicator	Porch Index of Communicative Ability	Stanford-Binet Intelligence Test		Sixteen (16) Personality Factor Questionnaire		
School and College Ability Tests	Purdue Pegboard	Wechsler Memory Scale	Wechsler Adult Intelligence Scale-R		Strong Interest Inventory		
Stanford Achievement Test	Raven Progressive Matrices		Wechsler Intelligence Scale for Children-R				
Tests of Achievement and Proficiency	Wesman Personnel Classification Test						
Wide Range Achievement Test							

Figure 1

Appendix N. Empirically Based Test Publisher Form: AGS Test Purchaser Qualification Form

AGS Test Purchaser Qualification Form

AGS requires all first-time individual test purchasers to furnish evidence of their qualifications to use tests. Test use should be consistent with sound professional practice, particularly those principles outlined in the 1985 APA, AERA, NCME publication *Standards for Educational and Psychological Testing*. Supply the information requested below, read carefully the "Principles of Effective Test Use," and sign the form to indicate acceptance of the responsibility for proper use of tests.

Identification Information (please complete all blanks)

Name_____ Position_____ Telephone ( )_____

Address: Street_____ City_____ State_____ Zip_____

Employer_____ Supervisor_____

Employer's Address: Street_____ City_____ State_____ Zip_____

A. Evidence of Appropriate Training in the Use of Tests

1. Your purpose for using tests (check as many as apply)

   ☐ Educational diagnosis/remediation ☐ Clinical diagnosis/classification ☐ Counseling ☐ Therapy ☐ Personnel selection ☐ Research
   ☐ Learning disabilities screening ☐ Other_____

2. Your level of training (check as many as apply)

   ☐ Bachelor's Degree: Year_____ Institution_____ Major field of study_____
   ☐ Master's Degree: Year_____ Institution_____ Major field of study_____
   ☐ Doctorate: Year_____ Institution_____ Major field of study_____

3. Your professional credentials (check as many as apply)

   ☐ Licensed in: Area_____ State_____ License #_____
   ☐ Member of professional organization(s): (AACD, AERA, APA, ASHA, CEC, NASP, etc.)_____
   ☐ Formally recognized professional competence (fellow, diplomate, special certificate): ☐ fellow ☐ diplomate ☐ other certification
   Organization_____

4. Your educational background (courses and other study)

   a. Courses (check each course completed and circle level at which course was completed.)
   U = Undergraduate, G = Graduate, O = Other (special course you have completed, workshop, in-service training, etc.)

	(circle)				(circle)		
☐ Basic tests and measurements	U	G	O	☐ Use of tests in counseling	U	G	O
☐ Descriptive statistics	U	G	O	☐ Career assessment	U	G	O
☐ Intelligence testing	U	G	O	☐ Neuropsychological assessment	U	G	O
☐ Speech, hearing, language assessment	U	G	O	☐ Other (list below)			
☐ Educational diagnostics	U	G	O	_____	U	G	O
☐ Assessment course in major field:	U	G	O	_____	U	G	O
_____				_____	U	G	O

   b. Other Study (check each type of program completed)

   ☐ Practicum in test administration and interpretation
   ☐ Internship (school psychology, counseling, etc.) type:_____

5. Your updating of professional knowledge and skills (check activities below and give an example)

   ☐ Attending workshops, seminars, conferences:_____
   ☐ Reading professional publications (journals, test manuals, etc.):_____

6. Your special competence
   (List one test which you use regularly that best illustrates your skill in test administration and interpretation.)

B. Evidence of Acceptance of the Responsibility for Sound Use of Tests

Principles of Effective Test Use

Sound, professional use of educational and psychological tests means that all test users must:

1. maintain the security of testing materials before and after testing.
2. avoid labeling students based on a single test score.
3. adhere strictly to the copyright law and under no circumstances photocopy or otherwise reproduce answer forms, test books, or manuals.
4. administer and score tests exactly as specified in the manual.
5. release results only to authorized persons and in a form in keeping with accepted principles of test interpretation.

Your signature indicates acceptance of and compliance with the above Principles.

Signature_____ Date_____

This form has been copyrighted by American Guidance Service.